REMARKS

Applicant respectfully requests reconsideration of the present application in view of the foregoing amendments and in view of the reasons that follow.

Applicants have amended pages 11 and 12 of the specification to overcome the objections raised by the Examiner.

Claims 1-15 were pending in the application. Claims 4-15, drawn to a non-elected invention, remain withdrawn from further consideration. Applicants reserve their right to file one or more divisional applications directed to any non-elected subject matter.

Applicants have revised claims 1-3 and added new claims 16-17. Upon entry of these amendments, claims 1-3 and 16-17 will be pending.

A detailed listing is presented, with an appropriate defined status identifier, of all claims that are or were in the application, irrespective of whether the claim(s) remain under examination. As noted, these amendments do not go beyond the original specification.

Objection to the Specification

The Examiner has noted that the terms "MARATHON-READY" and "PROMEGA" are trademarks and should be capitalized and accompanied with a generic terminology. With respect to the term ""MARATHON-READY," Applicants have amended the specification by deleting the sentence at page 11, lines 7-8 and introducing a new sentence that reads as follows:

> The 5'-RACE method was conducted following the protocol provided in the "Marathon-Ready™ cDNA Kit (mouse testis), which is commercially available from Clontech.

In addition, Applicants wish to point out that "PROMEGA" is a name of a manufacturer and, therefore, is not a tradename and does not have a generic terminology.

Furthermore, Applicants have assigned "SEQ ID NO:23" to the sequence on page 12, lines 17-18. Support for this amendment is found on page 13-14 of the enclosed Substitute Sequence Listing filed June 4, 2001 (Appendix A).

Rejection Under 35 U.S.C. § 112, Second Paragraph

The examiner alleges that the phrase "functionally equivalent" in claims 2 and 3, is indefinite and unclear because this phrase, as defined in the specification, is not indicative of the specific biological activities of the protein.

Contrary to the Examiner's allegation, the specification defines this phrase as a mutant protein having a "biochemical and/or biological activity" equivalent to the natural protein" and provides two examples of specific biological activities, namely, metal-binding activity and testicular cell differentiation-inducing activity (see specification at page 5, lines 7-9). However, to expedite prosecution, claims 2 and 3 have been amended to recite an isolated protein having testicular cell differentiation activity.

In light of the above remarks and foregoing amendment, Applicants submit that this rejection is moot and should be withdrawn.

Rejection Under 35 U.S.C. § 101

The examiner rejects claims 1-3 because these claims are drawn to non-statutory subject matter. To obviate this rejection, Applicants have followed the Examiner's suggestion by amending claims 1-3 to recite "an isolated protein..." Thus, this amendment renders the rejection moot.

Rejections Under 35 U.S.C. § 102

Finally, the examiner alleges that claims 2 and 3 lack novelty over Zon *et al.* because Zon teaches a DNA that encodes a mouse Tbc1 polypeptide "which likes tesmin, plays a role in testicular regulation." Moreover, the Examiner contends that Zon's Tbc1 polypeptide is "functionally equivalent to SEQ ID NO:4." In addition, the Examiner alleges that (i) "by sufficient addition, deletion and/or addition of SEQ ID NO:4..., one could obtain the protein of Zon *et al.*"; and (2) the sequence of Zon *et al.* "would hybridize under some conditions to the DNA that encodes SEQ ID NO:4." Applicants traverse the rejection.

In response, Applicants have amended claim 2 by deleting the phrase "have been replaced, deleted, and/or added, and which is functionally equivalent to the protein of claim 1" and by inserting the phrase "wherein said protein has testicular

found throughout the entire specification, particularly on page 5, lines 7-9.

To distinguish the protein of the cited art having similar biological activity, Applicants further revise claims 2 and 3 to define the claimed protein in terms of its homology or stringency of hybridization conditions. The description on page 6, lines 12-14 of the specification, serves as support for the specific homology. The stringent conditions for hybridization are specifically exemplified on page 6, lines 17-25 of the specification.

Applicants have introduced new claims 16-17 to recite an isolated protein having a higher homology. Support for this amendment is found in the specification on page 6, lines 12-16.

To further distinguish the claimed protein over the protein of Zon *et al.*, Applicants submit a BLAST alignment search result that compares the sequences of these two protein. As shown in Appendix B, there is no significant homology/similarity between the amino sequences of these two proteins or their corresponding nucleotide sequences that encode them. In fact, the length of both amino acid sequences or nucleotide sequences are also different. Thus, the claimed protein, Tesmin, is distinct from the mouse Tbc1 of Zon *et al.*

Based on the foregoing amendments and above remarks, Applicants believe that these rejections are rendered moot and should be withdrawn.

CONCLUSION

In view of the foregoing amendments and remarks, favorable reconsideration and allowance of this application are requested. An early notice in this regard is earnestly solicited. In the event that any issues remain, the Examiner is invited to contact the undersigned with any proposal to expedite prosecution.

Respectfully submitted,

Date Dec. 16, 2003

Stephen B. Maebius
Attorney for Applicants
Registration No. 35,264

FOLEY & LARDNER
Customer Number: 22428
22428
PATENT TRADEMARK OFFICE
Telephone: (202) 672-5569
Facsimile: (202) 672-5399

Should additional fees be necessary in connection with the filing of this paper, or if a petition for extension of time is required for timely acceptance of same, the Commissioner is hereby authorized to charge Deposit Account No. 19-0741 for any such fees; and applicant(s) hereby petition for any needed extension of time.

APPENDIX A

Docket No. 084335/0127

59

Receipt is hereby acknowledged of the following:

Applicants: Takashi Sugihara et al.
Serial No.: 09/743,237
Filing Date: January 5, 2001
For: TESTIS-SPECIFIC DIFFERENTIATION-REGULATORY FACTOR

1. Check No. 12418 for $240
2. Transmittal Letter to the U.S. Designated/Elected Office (2nd Submission) including; Copies of Form PCT/DO/EO/905 & 920; Statement to Support Filing and Submission in Accordance with 37 C.F.R. §§ 1.821-1.825; Substitute Sequence Listing (15 pages); and Associate Power of Attorney; Information Disclosure Statement and Form PTO-1449 with 1 attachment; Declaration and Power of Attorney; disk containing sequence listing; and Petition for Extension of Time for 1 month

Date Due: June 30, 2001
Date: June 4, 2001

Return to: SBM/mkm
Inspected by: SBM

FORM PTO-1390 (Modified) (REV 5-93) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: 084335/0127

		U.S. APPLICATION NO. (If known, see 37 C.F.R. 1.5) 09/743,237
INTERNATIONAL APPLICATION NO. PCT/JP99/03859	INTERNATIONAL FILING DATE July 16, 1999	PRIORITY DATE CLAIMED July 17, 1998

TITLE OF INVENTION
TESTIS-SPECIFIC DIFFERENTIATION-REGULATORY FACTOR

APPLICANT(S) FOR DO/EO/US
Takashi SUGIHARA, Renu WADHWA, Sunil C. KAUL, Youji MITSUI

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☐ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.
2. ☒ This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☐ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☐ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   - ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   - ☐ has been transmitted by the International Bureau.
   - ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)
6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   - ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   - ☐ have been transmitted by the International Bureau.
   - ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   - ☐ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern other document(s) or information included:

11. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information: Copies of Form PCT/DO/EO/905 & 920; Statement to Support Filing and Submission in Accordance with 37 C.F.R. §§ 1.821-1.825; Substitute Sequence Listing (15 pages); Associate Power of Attorney; and Petition for Extension of Time for 1 month

U.S. APPLICATION NO. (If known, see 37 C.F.R. 1.50	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
09/743,237	PCT/JP99/03859	084335/0127

17. ☒ The following fees are submitted:	CALCULATIONS	PTO USE ONLY
Basic National Fee (37 CFR 1.492(a)(1)-(5): Search Report has been prepared by the EPO or JPO.....$860.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482).....$690.00		
No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2).....$710.00		
Neither international preliminary examination fee (37 CFR 1.482) nor International search fee (37 CFR 1.445(a)(2)) paid to USPTO..... $1,000.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4).....$100.00		
ENTER APPROPRIATE BASIC FEE AMOUNT =		
Surcharge of $130.00 for furnishing the oath or declaration later than 20 Months from the earliest claimed priority date (37 CFR 1.492(e))		

Claims	Number Filed		Included in Basic Fee		Extra Claims		Rate		
Total Claims	15	-	20	=	0	x	$18.00		
Independent Claims	2	-	3	=	0	x	$80.00		
Multiple dependent claim(s) (if applicable)							$270.00		
TOTAL OF ABOVE CALCULATIONS =								$	
Reduction by ½ for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 CFR 1.9, 1.27, 1.28).									
SUBTOTAL =									
Processing fee of $130.00 for furnishing English translation later the 20 months from the earliest claimed priority date (37 CFR 1.492(f). +								$130.00	
TOTAL NATIONAL FEE =									
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). $40.00 per property +									
TOTAL FEES ENCLOSED =								$130.00	
							Amount to be: refunded	$	
							charged	$	

a. ☒ A check in the amount of $130.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. 19-0741 in the amount of $0.00 to the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 19-0741. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Foley & Lardner
3000 K Street, N.W., Suite 500
Washington, D.C. 20007-5109

SIGNATURE

NAME STEPHEN B. MAEBIUS

REGISTRATION NUMBER 35,264

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Atty. Docket No: 084335/0127

In re patent application of

SUGIHARA, TAKASHI *et al.*

Serial No. 09/743,237

Filed: January 5, 2001

For: TESTIS-SPECIFIC DIFFERENTIATION-REGULATORY FACTOR

STATEMENT TO SUPPORT FILING AND SUBMISSION IN ACCORDANCE WITH 37 C.F.R. §§ 1.821-1.825

Assistant Commissioner for Patents
Washington, D.C. 20231
Box SEQUENCE

Sir:

In connection with a Sequence Listing submitted concurrently herewith, the undersigned hereby states that:

1. the submission, filed herewith in accordance with 37 C.F.R. § 1.821(g), does not include new matter;

2. the content of the attached paper copy and the attached computer readable copy of the Sequence Listing, submitted in accordance with 37 C.F.R. § 1.821(c) and (e), respectively, are the same; and

3. all statements made herein of their own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United

States Code and that such willful false statements may jeopardize the validity of the application or any patent resulting therefrom.

Respectfully submitted,

May 31, 2001
Date

James A. Coburn

HARBOR CONSULTING
Intellectual Property Services
1500A Lafayette Road
Suite 262
Portsmouth, N.H.
800-318-3021

SEQUENCE LISTING

<110> SUGIHARA, TAKASHI
WADHWA, RENU
KAUL, SUNIL C.
MITSUI, YOUJI

<120> TESTIS-SPECIFIC DIFFERENTIATION-REGULATORY FACTOR

<130> 084335/0127

<140> 09/743,237
<141> 2001-01-05

<150> PCT/JP99/03859
<151> 1999-07-16

<160> 24

<170> PatentIn Ver. 2.1

<210> 1
<211> 2241
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (651)..(1535)

<400> 1
tatcctgtgg gttggccccg ggcagcaggc tctcagcagg ctcaggcacc acaagggata 60

cacagtgtgt gttcctggcc tgttggactt gtgactccac ccacctccgc cccagcaggg 120

ctagggatag aacccagggc cttttgcgtg ttctgcagat agtcttcagc ctggtagttt 180

ggggttggct gggagatttt tttttcttca caccaaagac ttccattatt gaggattttt 240

tcagttgatg atctcccccc tctgtaagat aagggacagt tctttaaacc tatgtagagt 300

tttgatgaat tctgcttctc aaccatattg ctaagctata tagcaattcc ttgaaattgc 360

tatataactt aggagaacct ctgattctcc tgcctctaca tcctgagtgc taggtgtaca 420

gggggaaatc attttggtga gactccgatg aactactgcc aggttcccaa ggcagcaagc 480

aagcaagaaa aagtgttgaa atcaaagaag caggtggtag tgtgccaggc ggcagccctg 540

aagacgcagc tttccaggcc cctctggctc aggaatcctg ttgcaagttc ccatcatccc 600

aggaggcaga ggaggcctcc agctgccctc ggaagaaaga ctccagcccc atg gtg 656
Met Val
1

att tgt cag ctg aaa gga ggc gcc cag atg ctc tgc ata gac aac tgt 704
Ile Cys Gln Leu Lys Gly Gly Ala Gln Met Leu Cys Ile Asp Asn Cys
5 10 15

```
ggc gcg agg gag ctc aaa gcg ctc cat ctg ctt cct cag tac gat gac   752
Gly Ala Arg Glu Leu Lys Ala Leu His Leu Leu Pro Gln Tyr Asp Asp
     20                  25                  30

cag agc agt ttc cct cag tca gag ctc cct aag cca atg aca act tta   800
Gln Ser Ser Phe Pro Gln Ser Glu Leu Pro Lys Pro Met Thr Thr Leu
 35                  40                  45                  50

gtg gga aga ctt ctg cca gta cca gcg aag tta aat ctc atc aca cag   848
Val Gly Arg Leu Leu Pro Val Pro Ala Lys Leu Asn Leu Ile Thr Gln
                 55                  60                  65

gtt gat aat gga gct ctc cca tca gct gtc aat ggg gct gcc ttt ccc   896
Val Asp Asn Gly Ala Leu Pro Ser Ala Val Asn Gly Ala Ala Phe Pro
             70                  75                  80

tct gga cct gct ctg caa ggg cca ccc aaa ata act ctg tct ggg tac   944
Ser Gly Pro Ala Leu Gln Gly Pro Pro Lys Ile Thr Leu Ser Gly Tyr
         85                  90                  95

tgt gac tgc ttc tcc agc ggg gac ttc tgc aac agc tgc agc tgc aac   992
Cys Asp Cys Phe Ser Ser Gly Asp Phe Cys Asn Ser Cys Ser Cys Asn
    100                 105                 110

aac ctg cgc cat gag ctc gag cgc ttc aaa gcc ata aag gcg tgt ctt  1040
Asn Leu Arg His Glu Leu Glu Arg Phe Lys Ala Ile Lys Ala Cys Leu
115                 120                 125                 130

gat aga aat cct gaa gct ttc caa cca aaa atg ggg aaa ggc cgt ctg  1088
Asp Arg Asn Pro Glu Ala Phe Gln Pro Lys Met Gly Lys Gly Arg Leu
                135                 140                 145

gga gct gct aaa ctt cga cac agc aaa ggg tgc aac tgt aag cgc tca  1136
Gly Ala Ala Lys Leu Arg His Ser Lys Gly Cys Asn Cys Lys Arg Ser
            150                 155                 160

ggc tgc ctg aag aac tac tgt gag tgc tat gag gcc aaa atc atg tgt  1184
Gly Cys Leu Lys Asn Tyr Cys Glu Cys Tyr Glu Ala Lys Ile Met Cys
        165                 170                 175

tct tcc att tgc aaa tgc att gct tgc aaa aac tat gaa gaa agt cca  1232
Ser Ser Ile Cys Lys Cys Ile Ala Cys Lys Asn Tyr Glu Glu Ser Pro
    180                 185                 190

gaa cga aaa atg ctg atg agc aca ccc cac tac atg gag cct ggg gac  1280
Glu Arg Lys Met Leu Met Ser Thr Pro His Tyr Met Glu Pro Gly Asp
195                 200                 205                 210

ttt gag agc agc cat tat ttg tcc cca gcc aag ttc tca gga cct cca  1328
Phe Glu Ser Ser His Tyr Leu Ser Pro Ala Lys Phe Ser Gly Pro Pro
                215                 220                 225

aaa ctg aga aaa aat agg cag gcc ttc tcc tgt atc tcc tgg gaa gta  1376
Lys Leu Arg Lys Asn Arg Gln Ala Phe Ser Cys Ile Ser Trp Glu Val
            230                 235                 240
```

```
gtg gag gcc aca tgt gcc tgc ctg ctg gcc cag ggt gag gaa gca gag   1424
Val Glu Ala Thr Cys Ala Cys Leu Leu Ala Gln Gly Glu Glu Ala Glu
        245                 250                 255

cag gag cac tgt tcc cca agc ttg gct gag cag atg atc ctg gag gag   1472
Gln Glu His Cys Ser Pro Ser Leu Ala Glu Gln Met Ile Leu Glu Glu
    260                 265                 270

ttt gga agg tgc ctg tcg cag att ctc cac atc gag ttc aag tcc aag   1520
Phe Gly Arg Cys Leu Ser Gln Ile Leu His Ile Glu Phe Lys Ser Lys
275                 280                 285                 290

ggg ctg aaa att gag tagcgtgcaa gctggtaaag gggaatgcct gtggcaagcc   1575
Gly Leu Lys Ile Glu
            295

tcagccctgg gaatctgcac cgaggaagct ggtgcccagg gaggagcaga ggccgcgcat 1635

catggccagg tcagctgtga ggtctgagtg atctgcatgg tactggccag cctactcaag 1695

gtatcctaaa gtgcaagcag gcagagccac cctggggatg gacactggcc ctcctgtccc 1755

tggggaggcc ctctggggac tccctgccct gcataaaaag agggtgattt tctacttgtt 1815

gttatgtgtt tgctttcaaa ttgcttagta gtacctccat tcaagttatt atgagccagc 1875

ctcaagttag agagctaggc tcttcttcag gtggactctg cccaaatcac atacaagtca 1935

ggtggccatc aggggttttt ccaggccagg cctgtgacag gagatatggg agggggtcg 1995

ggttagagct gggtttgttt ggatttttg cgttttttc ttcctgtatt tctgcttgaa 2055

gtgagaaaac ttgtctcctg tccaaccttt tctccataat tactgctgca cggtcgcctg 2115

ctgaccagtc acagtgacct cagacaccag aaggtgaggt ggcttattat gcccacactt 2175

tgtgttttgt tgtgagaata aacctttcca gactcccaaa aaaaaaaaaa aaaaaaaaaa 2235

aaaaaa                                                          2241


<210> 2
<211> 1861
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (271)..(1155)

<400> 2
ccagacgacg ccgctccgcc tcccgcctac agcgtgcacg tgttatcgtc gttacttccc 60

ggtgctcgcg ggcccgcgct gttgccgcta agcgcaggag tgcgcgtgat cccagttgaa 120

atcaaagaag caggtggtag tgtgccaggc ggcagccctg aagacgcagc tttccaggcc 180

cctctggctc aggaatcctg ttgcaagttc ccatcatccc aggaggcaga ggaggcctcc 240
```

```
agctgccctc ggaagaaaga ctccagcccc atg gtg att tgt cag ctg aaa gga  294
                                  Met Val Ile Cys Gln Leu Lys Gly
                                    1               5

ggc gcc cag atg ctc tgc ata gac aac tgt ggc gcg agg gag ctc aaa  342
Gly Ala Gln Met Leu Cys Ile Asp Asn Cys Gly Ala Arg Glu Leu Lys
     10                  15                  20

gcg ctc cat ctg ctt cct cag tac gat gac cag agc agt ttc cct cag  390
Ala Leu His Leu Leu Pro Gln Tyr Asp Asp Gln Ser Ser Phe Pro Gln
 25                  30                  35                  40

tca gag ctc cct aag cca atg aca act tta gtg gga aga ctt ctg cca  438
Ser Glu Leu Pro Lys Pro Met Thr Thr Leu Val Gly Arg Leu Leu Pro
                 45                  50                  55

gta cca gcg aag tta aat ctc atc aca cag gtt gat aat gga gct ctc  486
Val Pro Ala Lys Leu Asn Leu Ile Thr Gln Val Asp Asn Gly Ala Leu
             60                  65                  70

cca tca gct gtc aat ggg gct gcc ttt ccc tct gga cct gct ctg caa  534
Pro Ser Ala Val Asn Gly Ala Ala Phe Pro Ser Gly Pro Ala Leu Gln
         75                  80                  85

ggg cca ccc aaa ata act ctg tct ggg tac tgt gac tgc ttc tcc agc  582
Gly Pro Pro Lys Ile Thr Leu Ser Gly Tyr Cys Asp Cys Phe Ser Ser
     90                  95                  100

ggg gac ttc tgc aac agc tgc agc tgc aac aac ctg cgc cat gag ctc  630
Gly Asp Phe Cys Asn Ser Cys Ser Cys Asn Asn Leu Arg His Glu Leu
105                 110                 115                 120

gag cgc ttc aaa gcc ata aag gcg tgt ctt gat aga aat cct gaa gct  678
Glu Arg Phe Lys Ala Ile Lys Ala Cys Leu Asp Arg Asn Pro Glu Ala
                125                 130                 135

ttc caa cca aaa atg ggg aaa ggc cgt ctg gga gct gct aaa ctt cga  726
Phe Gln Pro Lys Met Gly Lys Gly Arg Leu Gly Ala Ala Lys Leu Arg
            140                 145                 150

cac agc aaa ggg tgc aac tgt aag cgc tca ggc tgc ctg aag aac tac  774
His Ser Lys Gly Cys Asn Cys Lys Arg Ser Gly Cys Leu Lys Asn Tyr
        155                 160                 165

tgt gag tgc tat gag gcc aaa atc atg tgt tct tcc att tgc aaa tgc  822
Cys Glu Cys Tyr Glu Ala Lys Ile Met Cys Ser Ser Ile Cys Lys Cys
    170                 175                 180

att gct tgc aaa aac tat gaa gaa agt cca gaa cga aaa atg ctg atg  870
Ile Ala Cys Lys Asn Tyr Glu Glu Ser Pro Glu Arg Lys Met Leu Met
185                 190                 195                 200

agc aca ccc cac tac atg gag cct ggg gac ttt gag agc agc cat tat  918
Ser Thr Pro His Tyr Met Glu Pro Gly Asp Phe Glu Ser Ser His Tyr
                205                 210                 215
```

```
ttg tcc cca gcc aag ttc tca gga cct cca aaa ctg aga aaa aat agg   966
Leu Ser Pro Ala Lys Phe Ser Gly Pro Pro Lys Leu Arg Lys Asn Arg
            220                 225                 230

cag gcc ttc tcc tgt atc tcc tgg gaa gta gtg gag gcc aca tgt gcc   1014
Gln Ala Phe Ser Cys Ile Ser Trp Glu Val Val Glu Ala Thr Cys Ala
        235                 240                 245

tgc ctg ctg gcc cag ggt gag gaa gca gag cag gag cac tgt tcc cca   1062
Cys Leu Leu Ala Gln Gly Glu Glu Ala Glu Gln Glu His Cys Ser Pro
    250                 255                 260

agc ttg gct gag cag atg atc ctg gag gag ttt gga agg tgc ctg tcg   1110
Ser Leu Ala Glu Gln Met Ile Leu Glu Glu Phe Gly Arg Cys Leu Ser
265                 270                 275                 280

cag att ctc cac atc gag ttc aag tcc aag ggg ctg aaa att gag       1155
Gln Ile Leu His Ile Glu Phe Lys Ser Lys Gly Leu Lys Ile Glu
                285                 290                 295

tagcgtgcaa gctggtaaag gggaatgcct gtggcaagcc tcagccctgg gaatctgcac 1215

cgaggaagct ggtgcccagg gaggagcaga ggccgcgcat catggccagg tcagctgtga 1275

ggtctgagtg atctgcatgg tactggccag cctactcaag gtatcctaaa gtgcaagcag 1335

gcagagccac cctggggatg gacactggcc ctcctgtccc tggggaggcc ctctggggac 1395

tccctgccct gcataaaaag agggtgattt tctacttgtt gttatgtgtt tgctttcaaa 1455

ttgcttagta gtacctccat tcaagttatt atgagccagc ctcaagttag agagctaggc 1515

tcttcttcag gtggactctg cccaaatcac atacaagtca ggtggccatc aggggttttt 1575

ccaggccagg cctgtgacag gagatatggg agggggtcg ggttagagct gggtttgttt 1635

ggatttttg cgttttttc ttcctgtatt tctgcttgaa gtgagaaaac ttgtctcctg 1695

tccaaccttt tctccataat tactgctgca cggtcgcctg ctgaccagtc acagtgacct 1755

cagacaccag aaggtgaggt ggcttattat gcccacactt tgtgttttgt tgtgagaata 1815

aacctttcca gactcccaaa aaaaaaaaaa aaaaaaaaaa aaaaaa                1861
```

<210> 3
<211> 2134
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (407)..(1303)

<400> 3

```
aattcggggt caaggcgaag ctcgcggggg gcgacagcga cggcggggag ctcctcgggg 60

agtaccccgg gatcccagag ctcagcgcgc tggaggacgt cgcgctcctg caggccccgc 120
```

```
agccgcccgc ctgcaacgtg cacttcctgt cctcgctgct acccgcgcac cgcagcccgc 180

gggtgttttg cccctggggc gcctgggtcc tgcgaaggag cctcccaccc gggcgtccgc 240

atgatcccag ttgaaatcaa ggtaagcagg tggtactact acaagtaata atccggaaga 300

agcaactttg cagaatcttc ttgctcagga atcctgttgc aagttcccat ggtcccagga 360

actagaggat gcctcctgct gttctcttaa gaaagattcc aaccca atg gtg ata    415
                                                 Met Val Ile
                                                   1

tgc caa ttg aaa ggg ggc aca caa atg cta tgt ata gac aat tct aga   463
Cys Gln Leu Lys Gly Gly Thr Gln Met Leu Cys Ile Asp Asn Ser Arg
      5                  10                  15

aca aga gaa cta aaa gca ctc cat ttg gtt cct cag tat caa gat caa   511
Thr Arg Glu Leu Lys Ala Leu His Leu Val Pro Gln Tyr Gln Asp Gln
 20                  25                  30                  35

aat aat tat cta cag tca gat gtc cct aaa cca atg act gct tta gta   559
Asn Asn Tyr Leu Gln Ser Asp Val Pro Lys Pro Met Thr Ala Leu Val
                 40                  45                  50

ggg aga ttt ttg cca gca tca aca aaa tta aat ctc att aca caa caa   607
Gly Arg Phe Leu Pro Ala Ser Thr Lys Leu Asn Leu Ile Thr Gln Gln
             55                  60                  65

ctt gag gga gcc tta cca tcg gta gtc aac ggg tct gct ttc ccc tcg   655
Leu Glu Gly Ala Leu Pro Ser Val Val Asn Gly Ser Ala Phe Pro Ser
         70                  75                  80

gga tca act ctt cca gga cca cca aaa ata act ttg gct ggg tac tgt   703
Gly Ser Thr Leu Pro Gly Pro Pro Lys Ile Thr Leu Ala Gly Tyr Cys
     85                  90                  95

gac tgc ttt gcc agt ggg gac ttt tgc aac aac tgc aat tgt aat aat   751
Asp Cys Phe Ala Ser Gly Asp Phe Cys Asn Asn Cys Asn Cys Asn Asn
100                 105                 110                 115

tgt tgc aac aac ttg cat cat gat att gaa cgg ttt aaa gcc att aag   799
Cys Cys Asn Asn Leu His His Asp Ile Glu Arg Phe Lys Ala Ile Lys
                120                 125                 130

gca tgt ctt ggt aga aat cca gaa gct ttc cag cca aaa att ggg aag   847
Ala Cys Leu Gly Arg Asn Pro Glu Ala Phe Gln Pro Lys Ile Gly Lys
            135                 140                 145

ggc caa ttg ggc aat gtc aag ccc cag cac aac aaa ggg tgc aac tgc   895
Gly Gln Leu Gly Asn Val Lys Pro Gln His Asn Lys Gly Cys Asn Cys
        150                 155                 160

agg agg tca ggc tgc ctg aag aat tac tgc gag tgc tat gag gcc caa   943
Arg Arg Ser Gly Cys Leu Lys Asn Tyr Cys Glu Cys Tyr Glu Ala Gln
    165                 170                 175
```

```
att atg tgt tct tct att tgc aaa tgc att ggt tgc aaa aat tat gaa   991
Ile Met Cys Ser Ser Ile Cys Lys Cys Ile Gly Cys Lys Asn Tyr Glu
180                 185                 190                 195

gaa agc cca gaa cga aag aca cta atg agc atg cca aac tac atg cag  1039
Glu Ser Pro Glu Arg Lys Thr Leu Met Ser Met Pro Asn Tyr Met Gln
                200                 205                 210

act gga ggt ttg gaa ggc agc cat tac ctg cca cca acg aaa ttt tca  1087
Thr Gly Gly Leu Glu Gly Ser His Tyr Leu Pro Pro Thr Lys Phe Ser
            215                 220                 225

gga ctt cca aga ttc agt cac gat agg cgg cct tcc tca tgc atc tcc  1135
Gly Leu Pro Arg Phe Ser His Asp Arg Arg Pro Ser Ser Cys Ile Ser
        230                 235                 240

tgg gag gtg gtg gag gcc aca tgc gcc tgc ctg ctt gct cag gga gaa  1183
Trp Glu Val Val Glu Ala Thr Cys Ala Cys Leu Leu Ala Gln Gly Glu
    245                 250                 255

gag gcc gag aaa gaa cac tgc tcc aag tgc ctg gca gag cag atg atc  1231
Glu Ala Glu Lys Glu His Cys Ser Lys Cys Leu Ala Glu Gln Met Ile
260                 265                 270                 275

ctg gag gaa ttt gga agg tgc tta tca cag att ctc cac act gag ttt  1279
Leu Glu Glu Phe Gly Arg Cys Leu Ser Gln Ile Leu His Thr Glu Phe
                280                 285                 290

aaa tct aag gga ttg aaa atg gag tagagtataa agtgtgaatg catgttgatt  1333
Lys Ser Lys Gly Leu Lys Met Glu
            295

ttgtcttagt ctagaaatct ctagtttaga aaggatgttt aggggaacat gaggctggct 1393

ctgcagcaac aaccaggctc ccctgcatcc ctgggcccag ggagtttact cagagctctc 1453

tgaagatgtg gcaacccatg ccccctttc tgaggaggtg catggcctga gcattgtttg 1513

tctggcccag aggagagagc ttgggttccc atagtcctgg gagagtgtct gcagggcggc 1573

ggagggcaga gcaggccctg cggagagctc actctggtcg actcttcctc tcagagaatg 1633

ttgctctgga ggctgctctg catgaaaacc ctaatggttt cttgtttgtt tttcaaatta 1693

tttagaaata agttctccgg atgggctgtt gtgataccac ttaaaatctc tagagaacta 1753

ctgaacacct aaagattttc tgtagcgtag atatttcccc agagacacgc gaactgtcag 1813

tctttcctaa ggcccccggg agacgcaggc aatggggcct cgcaggccag gcttgcacca 1873

gcatgtcttg agttagagga cttaaaatta tccagtttct tctgtgtttc tacttgaatt 1933

gtggaaaagc tctattatcc aattaacttc tccataatta ttgttgtaat attattattg 1993

tttgtaaaac atggttcaca taactagctt gtggaaacca gcaggtaaaa tgaattctta 2053
```

agttgacgct tttggttctg ttgtaaagca aagatgaata aaaatttcca atgtcgaaaa 2113

aaaaaaaaaa aaaaaaaaaa a 2134

<210> 4
<211> 295
<212> PRT
<213> Mus musculus

<400> 4

```
Met Val Ile Cys Gln Leu Lys Gly Gly Ala Gln Met Leu Cys Ile Asp
  1               5                  10                  15

Asn Cys Gly Ala Arg Glu Leu Lys Ala Leu His Leu Leu Pro Gln Tyr
             20                  25                  30

Asp Asp Gln Ser Ser Phe Pro Gln Ser Glu Leu Pro Lys Pro Met Thr
         35                  40                  45

Thr Leu Val Gly Arg Leu Leu Pro Val Pro Ala Lys Leu Asn Leu Ile
     50                  55                  60

Thr Gln Val Asp Asn Gly Ala Leu Pro Ser Ala Val Asn Gly Ala Ala
 65                  70                  75                  80

Phe Pro Ser Gly Pro Ala Leu Gln Gly Pro Pro Lys Ile Thr Leu Ser
                 85                  90                  95

Gly Tyr Cys Asp Cys Phe Ser Ser Gly Asp Phe Cys Asn Ser Cys Ser
            100                 105                 110

Cys Asn Asn Leu Arg His Glu Leu Glu Arg Phe Lys Ala Ile Lys Ala
        115                 120                 125

Cys Leu Asp Arg Asn Pro Glu Ala Phe Gln Pro Lys Met Gly Lys Gly
    130                 135                 140

Arg Leu Gly Ala Ala Lys Leu Arg His Ser Lys Gly Cys Asn Cys Lys
145                 150                 155                 160

Arg Ser Gly Cys Leu Lys Asn Tyr Cys Glu Cys Tyr Glu Ala Lys Ile
                165                 170                 175

Met Cys Ser Ser Ile Cys Lys Cys Ile Ala Cys Lys Asn Tyr Glu Glu
            180                 185                 190

Ser Pro Glu Arg Lys Met Leu Met Ser Thr Pro His Tyr Met Glu Pro
        195                 200                 205

Gly Asp Phe Glu Ser Ser His Tyr Leu Ser Pro Ala Lys Phe Ser Gly
    210                 215                 220

Pro Pro Lys Leu Arg Lys Asn Arg Gln Ala Phe Ser Cys Ile Ser Trp
225                 230                 235                 240

Glu Val Val Glu Ala Thr Cys Ala Cys Leu Leu Ala Gln Gly Glu Glu
                245                 250                 255
```

```
Ala Glu Gln Glu His Cys Ser Pro Ser Leu Ala Glu Gln Met Ile Leu
            260                 265                 270

Glu Glu Phe Gly Arg Cys Leu Ser Gln Ile Leu His Ile Glu Phe Lys
        275                 280                 285

Ser Lys Gly Leu Lys Ile Glu
    290                 295
```

<210> 5
<211> 299
<212> PRT
<213> Homo sapiens

<400> 5

```
Met Val Ile Cys Gln Leu Lys Gly Gly Thr Gln Met Leu Cys Ile Asp
  1               5                  10                  15

Asn Ser Arg Thr Arg Glu Leu Lys Ala Leu His Leu Val Pro Gln Tyr
            20                  25                  30

Gln Asp Gln Asn Asn Tyr Leu Gln Ser Asp Val Pro Lys Pro Met Thr
        35                  40                  45

Ala Leu Val Gly Arg Phe Leu Pro Ala Ser Thr Lys Leu Asn Leu Ile
    50                  55                  60

Thr Gln Gln Leu Glu Gly Ala Leu Pro Ser Val Val Asn Gly Ser Ala
65                  70                  75                  80

Phe Pro Ser Gly Ser Thr Leu Pro Gly Pro Pro Lys Ile Thr Leu Ala
                85                  90                  95

Gly Tyr Cys Asp Cys Phe Ala Ser Gly Asp Phe Cys Asn Asn Cys Asn
            100                 105                 110

Cys Asn Asn Cys Cys Asn Asn Leu His His Asp Ile Glu Arg Phe Lys
        115                 120                 125

Ala Ile Lys Ala Cys Leu Gly Arg Asn Pro Glu Ala Phe Gln Pro Lys
    130                 135                 140

Ile Gly Lys Gly Gln Leu Gly Asn Val Lys Pro Gln His Asn Lys Gly
145                 150                 155                 160

Cys Asn Cys Arg Arg Ser Gly Cys Leu Lys Asn Tyr Cys Glu Cys Tyr
                165                 170                 175

Glu Ala Gln Ile Met Cys Ser Ser Ile Cys Lys Cys Ile Gly Cys Lys
            180                 185                 190

Asn Tyr Glu Glu Ser Pro Glu Arg Lys Thr Leu Met Ser Met Pro Asn
        195                 200                 205

Tyr Met Gln Thr Gly Gly Leu Glu Gly Ser His Tyr Leu Pro Pro Thr
    210                 215                 220
```

Lys Phe Ser Gly Leu Pro Arg Phe Ser His Asp Arg Arg Pro Ser Ser
225 230 235 240

Cys Ile Ser Trp Glu Val Val Glu Ala Thr Cys Ala Cys Leu Leu Ala
245 250 255

Gln Gly Glu Glu Ala Glu Lys Glu His Cys Ser Lys Cys Leu Ala Glu
260 265 270

Gln Met Ile Leu Glu Glu Phe Gly Arg Cys Leu Ser Gln Ile Leu His
275 280 285

Thr Glu Phe Lys Ser Lys Gly Leu Lys Met Glu
290 295

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 6
ggagaatctg cgacaggcac 20

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 7
tccccagcca agttctcsgg 20

<210> 8
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 8
ttcattgacc tcaactacat g 21

<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 9
gtggcagtga tggcatggac 20

<210> 10
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 10
tatgggcgcc tcctttcagc tgacaaat 28

<210> 11
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 11
actgaggaag cagatggagc gctttgag 28

<210> 12
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 12
tgactgaggg aaactgctct ggtcat 26

<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 13
aacctgatgg ctggcttgat 20

<210> 14
<211> 20
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 14
tttttcttta ctttccttgg 20

<210> 15
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 15
ccgaagaagg gctccaagaa 20

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 16
tctccactca agacaagcct 20

<210> 17
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 17
tgggcccagg gagtttactc a 21

<210> 18
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 18
tctcccagga ctatgggaac ccaa 24

<210> 19
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 19
tcgaattcta tggtgatttg tcagctgaaa gga 33

<210> 20
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 20
gaattcgaat tcgcattccc ctttaccagc tt 32

<210> 21
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 21
accgtcgact gcctaaggtc ctgagaactt ggctgggga 39

<210> 22
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<400> 22
Cys Leu Ser Gln Ile Leu His Ile Glu Phe Lys Ser Lys Gly Leu Lys
1 5 10 15

Ile Glu

<210> 23
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<220>
<221> MOD_RES
<222> (2)
<223> Arbitrary amino acid

<220>
<221> MOD_RES
<222> (4)..(12)
<223> Arbitrary amino acid

<220>
<221> MOD_RES
<222> (14)
<223> Arbitrary amino acid

<400> 23
Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Cys
1 5 10 15

<210> 24
<211> 295
<212> PRT
<213> Mus sp.

<400> 24
Met Val Ile Cys Gln Leu Lys Gly Gly Ala Gln Met Leu Cys Ile Asp
1 5 10 15

Asn Cys Gly Ala Arg Glu Leu Lys Ala Leu His Leu Leu Pro Gln Tyr
20 25 30

Asp Asp Gln Ser Ser Phe Pro Gln Ser Glu Leu Pro Lys Pro Met Thr
35 40 45

Thr Leu Val Gly Arg Leu Leu Pro Val Pro Ala Lys Leu Asn Leu Ile
50 55 60

Thr Gln Val Asp Asn Gly Ala Leu Pro Ser Ala Val Asn Gly Ala Ala
65 70 75 80

Phe Pro Ser Gly Pro Ala Leu Gln Gly Pro Pro Lys Ile Thr Leu Ser
85 90 95

Gly Tyr Cys Asp Cys Phe Ser Ser Gly Asp Phe Cys Asn Ser Cys Ser
100 105 110

Cys Asn Asn Leu Arg His Glu Leu Glu Arg Phe Lys Ala Ile Lys Ala
115 120 125

Cys Leu Asp Arg Asn Pro Glu Ala Phe Gln Pro Lys Met Gly Lys Gly
130 135 140

```
Arg Leu Gly Ala Ala Lys Leu Arg His Ser Lys Gly Cys Asn Cys Lys
145                 150                 155                 160

Arg Ser Gly Cys Leu Lys Asn Tyr Cys Glu Cys Tyr Glu Ala Lys Ile
                165                 170                 175

Met Cys Ser Ser Ile Cys Lys Cys Ile Ala Cys Lys Asn Tyr Glu Glu
            180                 185                 190

Ser Pro Glu Arg Lys Met Leu Met Ser Thr Pro His Tyr Met Glu Pro
        195                 200                 205

Gly Asp Phe Glu Ser Ser His Tyr Leu Ser Pro Ala Lys Phe Ser Gly
    210                 215                 220

Pro Pro Lys Leu Arg Lys Asn Arg Gln Ala Phe Ser Cys Ile Ser Trp
225                 230                 235                 240

Glu Val Val Glu Ala Thr Cys Ala Cys Leu Leu Ala Gln Gly Glu Glu
                245                 250                 255

Ala Glu Gln Glu His Cys Ser Pro Ser Leu Ala Glu Gln Met Ile Leu
            260                 265                 270

Glu Glu Phe Gly Arg Cys Leu Ser Gln Ile Leu His Ile Glu Phe Lys
        275                 280                 285

Ser Lys Gly Leu Lys Ile Glu
    290                 295
```

APPENDIX B

NCBI **Blast 2 Sequences results**

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.6 [Apr-09-2003]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.00 wordsize: 11 Filter ☑

Sequence 1 gi 4581562 Homo sapiens tesmin mRNA, complete cds. **Length** 2134 Sug. hora

Sequence 2 gi 3206147 Sequence 1 from patent US 5700927 **Length** 4039 ZON

No significant similarity was found

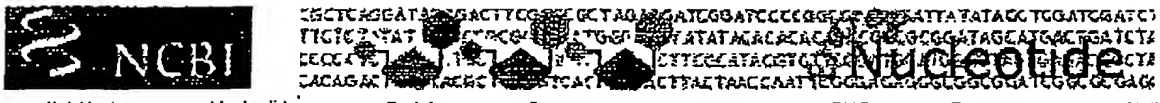
Blast 2 Sequences results

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.6 [Apr-09-2003]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 15 expect: 10.00 wordsize: 3 Filter

Sequence 1	gi 4758742	metallothionein-like 5, testis-specific (tesmin); tesmin [Homo sapiens] Sugihara	**Length** 299
Sequence 2	gi 3206311	Sequence 2 from patent US 5700927 ZON	**Length** 1141

No significant similarity was found



1: I86429. Sequence 1 from p...[gi:3206147] (Nucleotide)

```
LOCUS       I86429                  4039 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 1 from patent US 5700927.
ACCESSION   I86429
VERSION     I86429.1  GI:3206147
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4039)
  AUTHORS   Zon,L. and Richardson,P.
  TITLE     Tbc1 gene and uses thereof
  JOURNAL   Patent: US 5700927-A 1 23-DEC-1997;
FEATURES             Location/Qualifiers
     source          1..4039
                     /organism="unknown"
                     /mol_type="unassigned DNA"
BASE COUNT     1042 a   1096 c   1060 g    841 t
ORIGIN
        1 acctcggacc gaccgtctct tcaacccgca agagaagcag agactgtgga gtggaccccc
       61 accccagcag cgtggggcgg caagagcaca gccagctgag gaccatggac gccccgcccg
      121 cccaggccat cttctgggtg cccgggagca cctgcgcgcc ctagctcagc gcggagttct
      181 gagcgtgcca cctcactgga ttctgcccta agaagcacgt gaaatagatc acttcattta
      241 aagtaggaaa atagtagcac gtttttaccg gctcccaaga cctctcccag gatggaggca
      301 atcacattca cagcgaggaa gcatccgttt ccctaacgaa gtctccgtgg acttcggctt
      361 gcagctggtt gggtctcctt acccgtgcat tctctcacca ctatgcccat gctgccgtgg
      421 gtggtagccg aggtacgaag actcagtggc cagtgctcta aaaaggagcc taggacaaag
      481 caagtccgcc tttgggtttc gccctcggga ctgcggtgcg agcctgacct ggagaaaagt
      541 caaccatggg acccgctcat ctgttccagc atctttgagt gcaagcctca gcgtgtgcac
      601 aaactgattc acaacagtca cgacccgagc tactttgctt gcctcattaa ggaggatgct
      661 gcgcacaggc agagcctctg ctatgtgttt aaagcagacg atcaaacaaa agtgcctgag
      721 atcatcagct ccatccggca ggccgggaag attgcccgcc aggaagagct gcgttgcccc
      781 tccgagttcg acgatacctt cgccaaaaag ttcgaggtgc tcttctgtgg ccgggtgact
      841 gtggctcaca agaaggcccc acccgcactg attgacgagt gtatcgagaa gttcaaccat
      901 gtgagctgtg gtcgcagaac ggactgggaa gcgcccaccg ggcagccatc agcgcctggc
      961 cccaggccca tgcgcaaatc cttctcacag cctggactgc gcccgctggc cttcaggaag
     1021 gagttccagg acgctagcct ccgcagtagc acctttagct cctttgacaa tgacatagag
     1081 aaccacctca tcggtgggca caatgtggtt cagcccacag acatggagga gaaccgaact
     1141 atgctgttca cgattggccc atctgaagtt tacctcatca gtcctgacac caaaaagatt
     1201 gcactggaga aaaattttaa ggagatatcc ttttgctctc agggcatcag acatgtggac
     1261 cactttggat tcatctgccg agagtgctcg ggtggcggca gtggcggctt tcattttgtc
     1321 tgttacgtgt tccagtgcac aaatgaagct ctggttgacg agacatgat gactctgaag
     1381 caggctttca cggtagctgc ggtgcagcag acggctaagg caccagccca gctctgtgag
     1441 ggctgccctt tgcaaggcct gcacaagctc tgcgaaagga tagagggaat gaattcatct
     1501 aaaaccaaat tagaactcca gaagcacttg accacactga ccaacacagga gcaggccacc
     1561 atattcgagg aggttcagaa attgagacca agaaacgagc agcgagagaa tgaattaatt
     1621 atttcctttc tgaggtgctt atatgaagag aagcaaaaag agcacagcca cactggggcg
     1681 ccaaagcaga cactacaggt ggcagcagag aatattggga gtgacctgcc acccagtgct
```

```
1741 agccggttca ggttagattc gctgaagaac agagcaaaga ggtccttaac agagtcccta
1801 gagagcattc tgtcccgggg taataaagcc agaggcctgc aggaccattc cgccagtgtg
1861 gatctggaca gctccacttc tagtactcta agtaacacca gcaaagagct gtccatgggt
1921 gacaaggagg ccttccccgt ctctgagacc tccttcaagc tccttggctc ctcagatgac
1981 ctgtccagtg actcagaggg ccacattgca gaagagtctg ccctgttgtc accccagcag
2041 gcgttcagaa ggagagccaa caccctgagt catttcccag tagagtgccc tgcgcctcca
2101 gaacctgccc agagctctcc aggggtctct caaaggaaac tcatgcggta ccactccgtg
2161 agcacagaga cgcctcatga acgcaaggac tttgaatcca aagcaaacca cctgggtgac
2221 acagatggga ccccccgtgaa gaccccggcgg cactcgtgga gacagcagat attccttcga
2281 gtggccactc cacagaaggc ttgtgactcc ccgagcagat atgaagatta ttccgagctg
2341 ggagagcctcc ctccacgctc ccctttagaa ccggtgtgtg aggacggccc atttggccag
2401 tacaggaaga aaagaggaag acgtcacgcg agcttcgaga gctgtggaaa aaggccatct
2461 tgcagcagat cctgcctcgt caggatggag aaggagaatc agaagctaca agcctctgaa
2521 aacgatttgc tgaacaaacg cctcaagctt gactatgaag aaatcactcc gtgtcttaaa
2581 gaagtcacta cagtgtggga aaagatgctt agcactccag gaagatccaa aattaagttt
2641 gacatggaaa aagtgcactc agctgttggg caaggtgtgc cacgtcatca ccgaggtgag
2701 atctggaaat ttctagctga gcagttccac cttaaacacc catttcctag taaacagcag
2761 ccaaaggacg tgccctacaa agagctcctg aagaagctga cctcgcagca gcacgccatt
2821 ctcatcgacc tcgggcgaac ctttccaaca catccatact tctctgccca gcttggagca
2881 ggtcagctgt cactttacaa cattctgaag gcctactcgc ttctggacca ggaggttgga
2941 tactgccaag gtctcagctt tgtggcaggc attttgcttc ttcacatgag tgaggaagag
3001 gcgttcaaga tgctcaagtt cctgatgttt gacatggggc tgcggaaaca gtatcggcca
3061 gacatgatta ttttgcagat ccagatgtac cagctgtcac ggctcctcca cgattaccac
3121 cgagacctct acaaccacct ggaagagcac gagactggcc cccctacgta cgcggctccc
3181 tggtttctca ccgtgttcgc ctcacagttc ccactcggct ttgtagccag agtctttgat
3241 atgatcttcc ttcagggatc agaggtcata tttaaagtag ctttaagtct tttggggagc
3301 cataagccct tgattctaca gcatgagaac ctggaaacca tcgtggactt cataaagaac
3361 acactcccca acctgggcct ggtgcagatg gagaagacca tcagtcaggt gtttgagatg
3421 gacatcgcca agcagctcca ggcctatgag gtcgagtacc acgtcgtcca ggaggagctt
3481 attgagtcct cgcctctcag tgacaaccaa agaatggaga aattggagaa aaccaacagc
3541 acgttgcgca aacagaacct tgacctcctg gagcagttgc aggtggcaaa tgctaggatc
3601 caaagccttg aagccacggt agagaaactt cttaccagcg agagtaagct gaagcagcgt
3661 gcgctgaccc tggaggtgga gcgtcgccct gctgcagatg gtggaggagc tgcggaggca
3721 aagcgcccgg cccagcactc cagagccaga ctgcacccag ctggagccca caggcgattg
3781 accgctgcca gaagagactg tgcaccatta acactgtcca agccttaatc aagagagatg
3841 gaagtcagag gcagagaaga gagaacttct cagggaggaa actggctgac cagcctgcag
3901 atccttttga gctcagaact tgggattgga ggacaaaagt ctcagagtta ttgttgtttt
3961 tggttctaat ccgtcccctt tccagtcctg gttgttgtag ctttagatgg catggacatg
4021 aataaattac atttatggc
//
```

1: AAC19746. Sequence 2 from p...[gi:3206311] (protein) BLink, Links

```
LOCUS       I86593                1141 aa            linear   PAT 10-JUN-1998
DEFINITION  Sequence 2 from patent US 5700927.
ACCESSION   AAC19746
VERSION     AAC19746.1  GI:3206311
DBSOURCE    locus I86593 accession AAC19746.1
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (residues 1 to 1141)
  AUTHORS   Zon,L. and Richardson,P.
  TITLE     Tbc1 gene and uses thereof
  JOURNAL   Patent: US 5700927-A 2 23-DEC-1997;
FEATURES             Location/Qualifiers
     source          1..1141
                     /organism="unknown"
ORIGIN
        1 mpmlpwvvae vrrlsgqcsk keprtkqvrl wvspsglrce pdleksqpwd plicssifec
       61 kpqrvhklih nshdpsyfac likedaahrq slcyvfkadd qtkvpeiiss ircagkiarq
      121 eelrcpsefd dtfakkfevl fcgrvtvahk kappalidec iekfnhvscg rrtdweaptg
      181 qpsapgprpm rksfsqpglr slafrkefqd aslrsstfss fdndienhli gghnvvqptd
      241 meenrtmlft igpsevylis pdtkkialek nfkeisfcsq girhvdhfgf icrecsgggs
      301 ggfhfvcyvf qctnealvde immtlkqaft vaavqqtaka paqlcegcpl qglhklceri
      361 egmnssktkl elqkhlttlt nqeqatifee vqklrprneq reneliisfl rclyeekqke
      421 hshtgapkqt lqvaaenigs dlppsasrfr ldslknrakr slteslesil srgnkarglq
      481 dhsasvdlds stsstlsnts kelsmgdkea fpvsetsfkl lgssddlssd seghiaeesa
      541 llspqqafrr rantlshfpv ecpappepaq sspgvsqrkl nryhsvstet pherkdfesk
      601 anhlgdtdgt pvktrrhswr qqiflrvatp qkacdspsry edyselgelp prsplepvce
      661 dgpfgqyrkk rgrrhasfes cgkrpscsrs clvrmekenq klqasendll nkrlkldyee
      721 itpclkevtt vwekmlstpg rskikfdmek vhsavgqgvp rhhrgeiwkf laeqfhlkhp
      781 fpskqqpkdv pykellkklt sqqhailidl grtfpthpyf saqlgagqls lynilkaysl
      841 ldqevgycqg lsfvagilll hmseeeafkm lkflmfdmgl rkqyrpdmii lqlqmyqlsr
      901 llhdyhrdly nhleehetgp ptyaapwflt vfasqfplgf varvfdmifl qgsevifkva
      961 lsllgshkpl ilqhenleti vdfiknt1pn lglvqmekti sqvfemdiak qlqayeveyh
     1021 vvqeeliess plsdnqrmek lektnstlrk qnldlleqlq vanariqsle atvekllt se
     1081 sklkqraltl everrpaadg ggaeeakrpa qhsrarlhpa gahrrltaar rdcapltlsk
     1141 p
//
```

testis library was screened. As a result, a clone having an approximately 1.7 kb length was obtained.

Moreover, 5'-RACE was conducted to determine the 5' end sequence. In 5'-RACE, three antisense primers specific to the Tesmin gene, namely, SP1 (SEQ ID NO: 10), SP2 (SEQ ID NO: 11), and SP3 (SEQ ID NO: 12), and mouse testis-derived 5'-Marathon RACE cDNA were used. 5'-RACE method was conducted following the ~~"Marathon-ReadyTM~~ "Marathon-Ready™" cDNA kit (mouse testis)" (Clontech) protocol. The whole nucleotide sequence of Tesmin cDNA obtained is shown in SEQ ID NOs: 1 (2241 bp) and 2 (1861 bp). These two cDNAs are thought to be splicing variants arising from a difference in splicing at the point of transcription.

When the database was searched using these cDNA sequences, no sequence comprising a significant homology was found in the databank. These cDNAs encode the same protein comprising 295 amino acids (pI-7.64), and no significantly homologous proteins were found in the protein database as well.

Example 3: Cloning and sequencing of human Tesmin cDNA

Mouse Tesmin plasmid (a plasmid in which the Tesmin gene has been inserted into the pBluescript2 vector) was cleaved by SphI-SalI, and this 1.7 kb gene fragment was used as a probe to screen the cDNA library prepared by human testis mRNA. Hybridization was done using the "Rapid-hyb buffer" (Amersham LIFE SCIENCE) under the following conditions: (i) a prehybridization at 60°C for 30 min, (ii) addition of the labeled probe, and (iii) hybridization by incubating at 60°C for 2 hr. After that, washing is done three times within 2x SSC, 0.01% SDS for 20 min at room temperature, and next, three times within 1x SSC, 0.1% SDS, at 37°C for 20 min, followed by, two times within 1x SSC, 0.1% SDS, at 50°C for 20 min.

The nucleotide sequence of thus obtained human Tesmin cDNA is shown in SEQ ID NO: 3. Database search for the determined nucleotide sequence was done but there were no homologous sequences within the databank, similar to the mouse cDNA. The obtained human cDNA had four amino acids more than mouse Tesmin and encoded a

protein comprising 299 amino acids (pI-7.71). No significant homology was found in the protein database as well. However, as a result of amino acid sequence analysis by BLAST, the mouse and human Tesmins were found to be cysteine-rich proteins partially having the structure very similar to the metal-binding domain of the metallothionein family.

Metallothionein expression in the liver is induced by heavy metals, and metallothionein is known as a protein that neutralizes metallic poison. However, in the testis, the metallothionein gene is constantly expressed and is not induced by metals. Therefore, it was thought to play some vital roles other than metal binding in the testis. Recent findings showed that the estrogen receptor, which is a zinc-finger transcription factor and a receptor protein, and metallothionein conduct metal transfer *in vitro* (Cano-Gauci, D. and Sarkar, B. (1996) FEBS Lett 386 (1):1-4). Therefore, the metal-binding site of metallothionein is thought to play a vital role in the regulation of transcription factors. The "Cys-X-Cys-X-X-X-X-X-X-X-X-Cys-X-Cys (where X is an arbitrary amino acid)" sequence **(SEQ ID NO: 23)** having a cysteine structure in the amino acid sequence is thought to be vital for metal binding in the metallothionein family.

This cysteine structure (in mouse, from the 157th to the 171st positions, in human, from the 161st to the 175th positions) was conserved in Tesmin too. However, the metallothionein family members known up to now were relatively low-molecular comprising 60 to 70 amino acids, whereas Tesmin was comparatively longer (mouse: 295 amino acids, human: 299 amino acids). Domain search by PROSITE revealed that mouse Tesmin had a N-myristylation site and a casein kinase 2 phosphorylation site. Human Tesmin had a cAMP and cGMP-dependent kinase phosphorylation site, a protein kinase C phosphorylation site, a N-myristylation site, and a N-glycosylation site. Other than those, a cAMP and cGMP-dependent protein kinase phosphorylation site and a protein kinase phosphorylation site were also present.

Domain search by BLOCKS revealed sites common to mouse and human Tesmins. Namely, "high potential iron-sulfate protein"